DR. ANPING CHEN (Orcid ID : 0000-0003-2085-3863)

DR. DANIEL RICCIUTO (Orcid ID : 0000-0002-3668-3021)

DR. JINGFENG XIAO (Orcid ID : 0000-0002-0622-6903)

DR. LIANHONG GU (Orcid ID : 0000-0001-5756-8738)

Article type : Primary Research Articles

Moisture availability mediates the relationship between terrestrial gross primary production and solar-induced fluorescence: Insights from global scale variations

Anping Chen[1,*], Jiafu Mao[2,*], Daniel Ricciuto[2], Jingfeng Xiao[3], Christian Frankenberg[4,5], Xing Li[3], Peter E. Thornton[2], Lianhong Gu[2], Alan K. Knapp[1]

[1] Department of Biology and Graduate Degree Program in Ecology, Colorado State University, Fort Collins, CO, USA

[2] Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA

[3] Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH, USA

[4] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

[5] NASA Jet Propulsion Lab, California Institute of Technology, Pasadena, CA, USA

* Correspondence should be sent to anping.chen@colostate.edu (A.C.) and maoj@ornl.gov (J. M.)

Abstract

Effective use of solar-induced chlorophyll fluorescence (SIF) to estimate and monitor gross primary production (GPP) in terrestrial ecosystems requires a comprehensive understanding and quantification of the relationship between SIF and GPP. To date, this understanding is incomplete and somewhat controversial in the literature. Here we derived the GPP/SIF ratio from multiple data sources as a diagnostic metric to explore its global-scale patterns of spatial variation and potential climatic dependence. We found that the growing season GPP/SIF ratio varied substantially across global land surfaces, with the highest ratios consistently found in boreal regions. Spatial variation in GPP/SIF was strongly modulated by climate variables. The most striking pattern was a consistent decrease in GPP/SIF from cold-and-wet climates to hot-and-dry climates. We propose that the reduction in GPP/SIF with decreasing moisture availability may be related to stomatal responses to aridity. Furthermore, we show that GPP/SIF can be empirically modelled from climate variables using a machine learning (random forest) framework, which can improve the modelling of ecosystem production and quantify its uncertainty in global terrestrial biosphere models. Our results point to the need for targeted field and experimental studies to better understand the patterns observed and to improve the modelling of the relationship between SIF and GPP over broad scales.

1. Introduction

The total amount of carbon taken up by plants via photosynthesis, gross primary production (GPP), is a fundamental quantity for the global carbon cycle. The eddy covariance (EC) technique provides perhaps the best estimates of GPP at the ecosystem scale, despite numerous caveats (Goulden *et al*., 1996; Coops *et al*., 2007). With appropriate gap-filling, EC flux towers can continuously measure net ecosystem exchange (NEE), partitioned into carbon influx to the canopy (GPP) and outflux from the ecosystem (Respiration, R; Lasslop *et al*., 2010). However, EC towers are expensive to operate and their spatial extent is limited, challenging our ability to scale their measurements globally. Nonetheless, data-driven (or machine learning) methods have been used to upscale GPP from EC flux towers to regional, continental, and global scales using climate data and satellite observations (Xiao *et al*., 2010; Jung *et al*. 2011; Tramontana *et al.*, 2016). The upscaled GPP products are subject to the limitations of the original EC tower measurements (e.g., imbalanced spatial distribution with most sites in northern temperate areas; Schimel *et al*., 2015; Jung *et al.*, 2017; Wei *et al.*, 2017) as well as uncertainty from other input data, and incomplete representation of driving factors (Xiao *et al*., 2010). In addition to EC tower measurements and their scaled-up global products, remote sensing from space has been widely used for GPP monitoring at regional to global scales (Schimel *et al*., 2015; Huemmrich *et al.*, 2019). Compared to EC-based GPP estimates, remote sensing provides efficient GPP monitoring across all the Earth's land surface without being restricted by site accessibility (Running *et al*., 2000). A wide range of remote sensing approaches based on vegetation indices, light use efficiency (LUE) models and solar-induced chlorophyll fluorescence (SIF), have been used for GPP estimation and monitoring (Hilker *et al*., 2008; Sun *et al*., 2018; Xiao *et al*., 2019a). For example, vegetation greenness indices have been widely used to approximate GPP because of their close relationship with photosynthetic capacity (Tucker et al., 1986) and these have revealed a greening trend of the global land since the 1980s (Mao *et al*., 2016; Piao *et al*., 2020; Zhu *et al*., 2016). However, informing GPP from greenness indices is subject to biases from radiometric saturation (Tan *et al.*, 2013), contamination by atmospheric scattering, cloud, and background reflectance (Kaufman *et al.*, 1993; Gutman *et al.*, 1987), and the failure to represent transient environmental stresses (Guanter *et al*., 2014b). The first global operational GPP product, Moderate Resolution Imaging Spectroradiometer (MODIS) GPP, is based on the LUE method driven by

remotely sensed data (Running *et al*. 2004). However, the LUE method is subject to uncertainties in remote sensing data, model structure, and parameterization (Xiao *et al*., 2019a).

More recently, SIF has emerged as a promising approach for monitoring seasonal variations of GPP from regional to global scales (Frankenberg *et al.*, 2011; Guanter *et al.*, 2012; Jeong *et al*., 2017; Joiner *et al.*, 2011). The emergence of the SIF data stream from satellite passive measurements may overcome many drawbacks of traditional greenness-based satellite products, leading to vast improvements in estimating GPP spatially and temporally because of the strong relationship between GPP and SIF. For example, by using SIF we can avoid extrapolation of GPP using greenness indices that may fail to capture transient stresses. Mechanistic studies suggest that the relationship between SIF and GPP can be described with a simple conceptual relationship: $GPP = SIF \times \frac{LUE}{\Phi_F \times \varepsilon}$, where Φ_F is the fluorescence yield, and ε is the probability of an emitted photon being detected from above the canopy (Frankenberg and Berry, 2018). This implied strong correlation between SIF and GPP has been demonstrated by many empirical studies from single leaves to ecosystems, and over daily to yearly time scales (Damm *et al.*, 2010; Campbell *et al.*, 2007; Yang *et al.*, 2015; Zhang *et al.*, 2016a, 2016b; Lee *et al.*, 2015; Verrelst *et al.*, 2015, 2016; Rossini *et al.*, 2015, 2016). Interestingly, the relationship between leaf-level SIF and GPP is often nonlinear at short time scales (i.e., hourly) but becomes increasingly linear at longer time scales and broader spatial scales (Li *et al.*, 2018a; Damm *et al.*, 2015; Jung *et al.*, 2011; Yang *et al.*,2017; Gu *et al.*, 2019; Frankenberg *et al.*, 2014; Guanter *et al.*, 2012). This cross-scale shift in GPP-SIF relationships is usually attributed to the mean-field effect of larger scales that cancels out nonlinear terms (Thum *et al.*, 2017). For the purpose of investigating the carbon cycle at these larger spatial and temporal scales, global change biologists thus usually assume that SIF and GPP are linearly related (Frankenberg *et al.*, 2011; MacBean *et al.*, 2018; Parazoo *et al*., 2014).

What is unresolved is how this linear GPP-SIF relationship varies across different regions and vegetation types (Gu *et al.*, 2019; MacBean *et al.*, 2018; Xiao *et al.*, 2019b; Li *et al.*, 2018a; Zhang *et al.*, 2018). For example, by comparing SIF data from the Airborne Prism EXperiment (APEX) sensor and GPP data from EC towers, Damm *et al*. (2015) reported variable GPP-SIF relationships across vegetation types, with the strongest relationships in croplands and the weakest in mixed temperate

forest. More recent studies (e.g., Sun *et al.* 2017, Li *et al.* 2018a) found consistent linear relationships across biomes between SIF from the Orbiting Carbon Observatory-2 (OCO-2) and GPP from EC towers. These contrasting GPP-SIF findings may arise from the differing satellite products, or the scales and methods of the investigations (Yang *et al.*, 2015; Sun *et al.* 2017, Li *et al.* 2018a, 2018b). In particular, different SIF products may have different systematic biases that need to be recognized when interpreting the GPP-SIF relationship across vegetation types. To resolve the controversy, the GPP-SIF relationship needs to be assessed across data sources and evaluated in a more general environmental framework, particularly across climatic space (Thum *et al.* 2017; Smith *et al.*, 2018; Qiu *et al.*, 2019). Such assessments will allow us to address key issues such as how the GPP-SIF relationship changes along latitudinal gradients and which climatic variables play dominant roles in regulating variations in the relationship. Exploring general patterns of variation in the GPP-SIF relationship across climates, regions, and vegetation types, and mechanistic explanations of such patterns is not only important to better understand and predict this relationship, but also can inform and improve the modelling of GPP under current and future climates (Lee *et al.*, 2015).

Here, we assembled multiple satellite derived SIF data sets and both satellite and ground-based GPP data, along with corresponding climate (temperature, precipitation, and shortwave radiation) variables, to investigate how the GPP-SIF relationship may change across the global land surface. Our primary goal was to uncover potential environmental controls on the GPP-SIF relationship and to generate hypotheses regarding the underlying mechanisms. In doing so, we used four different SIF products in combination with two independent GPP datasets. Using the GPP/SIF ratio as a diagnostic metric of the GPP-SIF relationship, we examined its variation across both geographic and climatic spaces. To further quantify climatic controls on the spatial variation of GPP/SIF, we also developed a machine learning model (random forest) of GPP/SIF based on the above climatic variables.

2. Data and Methods

Here we used the GPP/SIF ratio as a metric to diagnose the dynamics of the GPP-SIF relationship. Unlike other metrics, such as GPP/temperature or GPP/precipitation for which temperature or precipitation is an environmental factor driving variations in GPP (e.g., Piao *et al*.,

2013), photosynthetic carbon assimilation and fluorescence re-emission are parallel processes during photosynthesis. Thus, the GPP/SIF ratio can be interpreted as a diagnostic metric sensitive to how photon quanta are allocated to different photochemical processes (Gu *et al.*, 2019; Lee *et al.*, 2015).

2.1 Satellite-based SIF products

We used four gridded and contiguous SIF products derived from different satellite platforms/instruments, including Global Ozone Monitoring Experiment-2 (GOME-2; Joiner *et al.*, 2013), Greenhouse Gases Observing Satellite (GOSAT; Frankenberg *et al.*, 2011), a finer-resolution global SIF product derived from Orbiting Carbon Observatory 2 (GOSIF; Li & Xiao, 2019), and Tropospheric Monitoring Instrument (TROPOMI; Köhler *et al.*, 2018). These SIF products have different spatial and temporal resolutions as well as different temporal coverages. All the SIF data have been corrected for day length difference and were linearly interpolated to a common 1° by 1° spatial resolution.

The GOME-2 SIF was retrieved from sensors onboard the EUMETSAT's Meteorological Operational Satellite-A (MetOp-A) for the period 2007–2018 (Köhler *et al.*, 2015a). These SIF retrievals were obtained from the 715–758 nm spectral window. GOME-2 SIF is available globally at 0.5° by 0.5° spatial resolution and daily temporal resolution. The GOSAT SIF was produced by the Japan Aerospace Exploration Agency (JAXA) at the spatial resolution of 3˚ by 3˚ (Frankenberg *et al.*, 2018). The GOSIF was a SIF product derived from multiple data sources, and produced with high spatial resolution (0.05˚ by 0.05˚) and 8-day temporal resolution for the period 2000–2018 (Li & Xiao, 2019). This SIF product was produced by training the available footprints (1.3 km × 2.25 km) of the OCO-2 data and MODIS-derived vegetation indices and reanalysis-based climatic variables with the Cubist regression tree model (Li & Xiao, 2019). In this way, it extended the current short data record of OCO-2 SIF (since September 2014) to a much longer period (since March 2000) and sparse footprints to global coverage (Li & Xiao, 2019). TROPOMI SIF was derived from a spectrometer onboard the Sentinel-5 Precursor satellite launched in October 2017 by the European Space Agency (Köhler *et al.*, 2018). The TROPOMI-based gridded SIF data have greatly improved temporal (~1 day) and spatial resolution (e.g., 0.2° by 0.2°) compared to GOME-2 and GOSAT SIF data. For all SIF data datasets, we used the daily corrected values provided by the respective products. The daily

correction transforms instantaneous SIF acquisition at the overpass time to a daily average (Frankenberg *et al*., 2011), and thus ensures comparability among different data sources.

2.2 Flux tower- and satellite-based GPP products

Two GPP datasets were used, one obtained from the MOD17A2 model (MODIS GPP; Zhao and Running, 2010) and one from EC tower upscaling FLUXCOM (Tramontana *et al.*, 2016). The MODIS GPP product was derived from an empirical LUE model driven by MODIS measurements of fPAR (fraction of Photosynthetically Active Radiation) and LAI (Leaf Area Index) in conjunction with meteorological reanalysis datasets (Zhao and Running, 2010). Monthly gridded MODIS GPP was provided at 0.05° by 0.05° spatial resolution from 2000 to 2018 across the globe. We aggregated this high-resolution dataset to 1° by 1° grids to maintain consistency with the SIF products. The FLUXCOM GPP product was derived from training three machine learning algorithms (model trees ensemble, artificial neural networks, and multivariate adaptive regression splines and random forests) on in-situ data from 224 FLUXNET sites using meteorological measurements and satellite data as inputs (Tramontana *et al.*, 2016). The FLUXCOM GPP is available for the 1982-2016 period, with a 0.5° by 0.5° spatial resolution and a monthly temporal resolution.

We also retrieved site-level flux tower GPP data to test the robustness of results from gridded datasets. We only selected sites (1) where the local vegetation type is representative of the dominant vegetation type of the corresponding 1° by 1° pixel based on the MODIS land cover map; (2) that have an overlapping observation period with the SIF data. Applying these selection criteria led to a subset of 53 sites for GOME2 and GOSIF and 49 sites for GOSAT (Figure S1). To derive site-level GPP/SIF values, we extracted SIF values for each site from its corresponding 1° by 1° grid cell. Note that due to the limited number of flux tower sites that met our criteria, we used all the available monthly data during the growing season in this analysis. Min-max scaling (see below) was applied separately to the site-month GPP and SIF values before calculating the GPP/SIF.

2.3 Climate and soil moisture data

Gridded monthly temperature and precipitation data were obtained from the Climatic Research Unit (CRU) (Harris *et al.*, 2014) v4.0.1 data set (http://www. cru.uea.ac.uk/data/), produced by the University of East Anglia in UK. These climatic data at 0.5° by 0.5° spatial resolution were

interpolated from meteorological station observations (Harris *et al.*, 2014). The latest version of the CRU climatic data is available for the period 1901-2018, covering the duration of all GPP and SIF products. Downward shortwave radiation (SWR) data were derived from the CRUNCEP V8 (Viovy, 2016), which is a combination of CRU monthly climatology and 6-hourly NCEP reanalysis meteorological data. In addition, soil water is the direct source of water supply for plant photosynthesis; and soil water content (SWC), which integrates the signal of both temperature and precipitation, influences both GPP and SIF more directly than those climatic factors. We used SWC data from the European Space Agency (ESA)'s Soil Moisture Ocean Salinity (SMOS) Earth Explorer mission processed with the INRA-CESBIO algorithm (SMOS-IC), which is based on the same SMOS Level 2 SM algorithm but has more independence from auxiliary data (Fernandez-Moran et al., 2017). SMOS-IC version 105 is at a spatial resolution of 25km by 25km, spanning from 1982 to 2018 (Fernandez-Moran et al., 2017), and its uncertainty at global scale has been evaluated. The quality of SMOS-IC data is assured by direct comparison with in situ soil moisture networks (Al Bitar et al., 2017; Jackson et al., 2012). Finally, we grouped the global land grids into three climate zones based on the Köppen–Geiger classification system (Peel et al., 2007). These three climate zones are: the tropical region defined as the Köppen-Geiger A (tropical) climate group; the temperate region as the Köppen-Geiger B (dry) and C (mild temperate) climate group; and the boreal region as the Köppen-Geiger D (continental) and E (polar) climate group.

2.4 Analyses

Because of inherently low signal-to-noise ratio of the SIF measurements and uncertain bias correlation in the non-growing season, our analyses focused on the growing season. Here grid-specific growing season was derived using the methods described in Zhu *et al*. (2016), and we calculated multi-year mean GPP/SIF ratio during this defined period. To facilitate inter-comparability among different data sets, we first normalized each GPP and SIF product by min-max scaling (i.e., making GPP or SIF values across the globe within the range of [0,1]). Results derived from non-normalized products are also presented to check if the normalization will profoundly change the geographic pattern (Figure S2). We then calculated the GPP/SIF ratio (i.e., GPP divided by SIF) using all available combinations of GPP and SIF data with overlapping periods. Hence, TROPOMI SIF (since March 2018) were only combined with MODIS GPP since currently available FLUXCOM GPP

dataset ended at December 2016. We thus obtained seven combinations of GPP/SIF patterns and calculated the multi-product mean for further analyses. The different time periods of SIF products can be expected to have minor impacts on our study, as we focus solely on the spatial variations. Furthermore, to investigate if different growing season lengths may change the observed patterns, we also conducted the same analyses with data from only the GPP-peak month (i.e., the month with the maximum GPP value over a year).

In assessing the spatial pattern of climatic control on GPP/SIF, we performed partial correlation analyses of GPP/SIF against temperature, precipitation, and shortwave radiation with 9° by 9° degree moving windows. We also used a machine learning method (random forest (RF), Breiman, 2001) to explain spatial patterns of the GPP/SIF ratio using temperature, precipitation, and shortwave radiation as explanatory variables. Machine learning is a modeling solution that differs fundamentally from simple regression models and complex simulation models. It is particularly powerful for dealing with large-scale multivariable data in which complex relationships exist among predictors (Chen *et al*., 2019; Reichstein *et al*., 2019; Tramontana *et al*., 2016). The RF model was trained on all vegetated grid cells, and then used to predict the spatial pattern of the GPP/SIF ratio from the same set of explanatory variables. The performance of this model was evaluated by comparing the predicted values against the trained values. The RF model permits an assessment of the importance of explanatory variables, through the increase in mean squared error of the RF predictions after permuting each explanatory variable (Gregorutti *et al.*, 2017).

3. Results

3.1 Spatial distribution of the GPP/SIF ratio

The spatial distribution of the GPP/SIF ratio, based on the average of different GPP and SIF product combinations, showed a striking pattern of high GPP/SIF ratio in the northern high latitudes (mostly ~60 °N), and in the tropical rainforests in the Amazon, Congo Basin, and Southeast Asia (Figure 1a). By contrast, dry regions, including central Asia, Southwest America, Australia, and Southwest Africa, had lower GPP/SIF ratios. As a result, along the latitudinal gradient, the GPP/SIF ratio followed a bi-modal hump-shaped curve (Figure 1a left panel). However, for both GPP and SIF

individually, their latitudinal distributions displayed unimodal hump-shaped patterns, with peaks around the equator (Supplementary Figure S2). In addition, GPP/SIF was more strongly related with GPP (R^2=0.59) than with SIF (R^2=0.28) (Figures 2e, 2f). Finally, the largest uncertainties of GPP/SIF occurred in the arid and boreal regions (Figure 1b), sourced from the strong discrepancy of SIF values across different satellite products. In arid regions, very low SIF values but large uncertainties indicate less reliable relationships between GPP and SIF.

The pattern of high GPP/SIF ratios in the northern high latitudes was generally consistent across different GPP and SIF product combinations (Supplementary Figure S2) and that from GPP-peak month (Supplementary Figure S3). Note that GPP and SIF values were not normalized in Figures S2, indicating that normalization does not change the overall geographic distribution pattern of GPP/SIF ratios. Similarly, the consistent geographic distributions of GPP/SIF between the entire growing season and the GPP-peak month indicates that growing-season length variation had little influence on GPP/SIF patterns. However, a number of different GPP and SIF combinations showed notable differences (Supplementary Figure S2). For example, in FLUXCOM-GOSIF, MODIS-GOSIF, and MODIS-TROPOMI, the GPP/SIF ratio for tropical rainforests was significantly lower than that for northern high latitudes (Figure S2c, S2f, S2g). Drylands, particularly those in Australia, also showed unusually high GPP/SIF ratios for GOSIF- or TROPOMI-derived results (Figure S2c, S2f, S2g). Furthermore, non-normalized data showed more uneven patterns along latitudinal gradient (Figure S2).

3.2 GPP/SIF ratios across the climate space

In climate space, the GPP/SIF ratio decreased from wet-and-cold to dry-and-hot regions for the average of all GPP and SIF product combinations (Figure 3), reflecting a gradient of decreasing moisture availability. It appeared that the GPP/SIF ratio generally increased with precipitation at a given temperature but varied less with temperature when precipitation was fixed (Figure 3). The decreasing GPP/SIF trend from wet-and-cold to dry-and-hot was also confirmed by examining individual product combinations of GPP and SIF (Supplementary Figure S4), except for the combination of FLUXCOM-GOSIF which showed no clear pattern (Figure S4c). In addition, because absorbed photosynthetically active radiation (APAR) is directly related to vegetation biogeophysics and photosynthetic processes, APAR could also drive the GPP/SIF variations. Indeed, the GPP/SIF

ratio showed an even clearer pattern in the APAR-precipitation space, increasing along both the APAR and precipitation gradients (Figure 3).

The dominant role of precipitation in controlling GPP/SIF ratios was further indicated by the spatial distribution of the partial correlation coefficients of GPP/SIF against temperature, precipitation, and shortwave radiation (Supplementary Figure S5). Partial correlation analyses with moving windows showed that the temperature correlation with GPP/SIF varied across different GPP and SIF product combinations, with both positive and negative correlation coefficients. Overall, the partial correlation between GPP/SIF ratios and temperature tended to be significantly positive ($P < 0.05$) in boreal regions along the ~60°N latitude, but was significantly negative ($P < 0.05$) in many temperate regions (Figure S5a-S5g). For the tropics, the results were mixed depending on the products used (Figure S5a-S5g). Similarly, mixed results could be found for the correlations with shortwave radiation (Figure S5o-S5u). However, for precipitation, its partial correlation with GPP/SIF ratios was overwhelmingly positive ($P < 0.05$) except for some boreal and tropical regions (Figure S5h-S5n). In addition, SWC was able to explain 47% of the spatial variation in GPP/SIF ratios (Figure S6), confirming the primary role of moisture availability in controlling GPP/SIF ratios as revealed in Figures 3 and S5.

To further assess this pattern, we conducted the same analyses using site-level flux tower GPP and the corresponding original satellite SIF retrievals (Figure 4). Although not as strong as the pattern derived from gridded data (Figures 3 and S4), results from site-level analyses confirmed the trend of decreasing monthly GPP/SIF with climate changing from cold-and-wet climates to hot-and-dry, especially for GOSIF (Figure 4). For GOME-2 and GOSAT products, site-level GPP/SIF was primarily controlled by precipitation.

3.3 Machine learning modelling of GPP/SIF ratios by climatic variables

Using a random forest (RF) machine learning method to model the GPP/SIF ratio with the three climatic variables: temperature (TMP), precipitation (PRE), and SWR, the three variables explained about 88% of the spatial variations in the GPP/SIF ratio, with PRE the most important explanatory variable, followed by SWR and TMP (Figure 5). The high explanatory power of the three climatic variables with an RF algorithm was also consistent with that based on individual GPP and SIF product

combinations (Supplementary Figure S7), where R^2 values varied between 0.72 and 0.88. The highest R^2 value was found for the FLUXCOM-GOME-2 combination, while the lowest was observed for both FLUXCOM-GOSIF and MODIS-GOSIF combinations. The ranking of the relative importance by the three climatic variables, PRE>SWR>TMP, was also consistently found across the different product combinations, except for FLUXCOM-GOSIF where SWR>TMP>PRE (Figure S7). Interestingly, the relative importance of different climatic variables varied across different climate zones (Figure 5). Consistent with the global-scale result, the spatial variation of the GPP/SIF ratio in temperate regions (including the arid region) was primarily dominated by PRE. However, in boreal and tropical regions, it was dominated by TMP and SWR, respectively. This regional divergence of dominant factors for the GPP/SIF ratio is consistent with proposed regional climate constraints of plant growth (Nemani et al., 2003).

4. Discussion

The GPP/SIF ratio represents a useful metric for understanding the relationship between the GPP and SIF signals, with the latter being increasingly used for understanding photosynthetic carbon assimilation at regional to global scales (Xiao et al., 2019a). Here we conducted a global-scale investigation on the distribution of the GPP/SIF ratio, and examined its covariation with climatic variables. Li *et al*. (2018a) also examined the relationship between GPP and SIF, but at daily or sub-daily scales using data from OCO-2 and tower GPP for 64 sites across the globe. They found largely invariant GPP-SIF relationships across different biomes. Our research builds on this understanding of the GPP-SIF relationships by integrating multiple datasets including both gridded and site-level data, and by examining the climatic drivers. In contrast to the findings of Li *et al*. (2018a), we found that the growing season GPP/SIF ratio varied greatly across the global land surface, with high values mostly found in boreal regions and tropical rainforests. This large spatial variation of GPP/SIF was also observed with different GPP and SIF product combinations, and across the entire growing season or for the GPP-peak month, despite divergent spatial patterns with different products. The finding thus suggests that the GPP-SIF relationship is not constant, but instead is modified by environmental influences. In contrast to Li *et al.* (2018a) which employed data from 64 site observations at the daily or sub-daily level, here we used gridded data integrative of the full growing season covering all the

global land. The different spatial and temporal scales may explain the divergent findings between these two studies.

Currently available GPP and SIF products are both known to have large systematic biases, particularly when the resolution is coarse (Frankenberg et al., 2014). Such biases could have affected the observed GPP-SIF relationship. However, some consistent GPP/SIF geographic patterns with different GPP and SIF product combinations and with different time scales (the entire growing season versus the peak-growth month), such as the high GPP/SIF ratio in the northern high latitudes, were not likely to be impacted by systematic data biases. On the other hand, further investigations are needed to understand some divergent patterns of the GPP/SIF ratio across different satellite products. For instance, drylands were found to have high GPP/SIF ratios with GOSIF- or TROPOMI products. Note from the GPP/SIF scale in Figure S2, the SIF signal seems to be smaller in GOSIF or TROPOMI products than that of the other two SIF products. It is thus possible that the SIF signal in drylands from GOSIF or TROPOMI products is close to minimum, resulting in high GPP/SIF ratios estimated with the two SIF products in these drylands.

More importantly, despite substantial variation across geographic space and among different satellite products, we found that the GPP/SIF ratio showed strong patterns along climatic gradients. Particularly, there was a decreasing trend in growing season GPP/SIF from cold-and-wet climates to hot-and-dry (i.e., along a gradient of moisture availability). The robustness of our finding was generally corroborated by different sets of GPP and SIF data, by results from the peak-growth month, and by site-level data analyses. Li *et al*. (2018a) suggested APAR as the primary factor driving GPP/SIF patterns. Indeed, the dominant role of APAR in shaping the variation of GPP/SIF ratios was also confirmed by our study (Figure 3). Nonetheless, APAR represents photosynthetically active radiation and is directly related to vegetation biogeophysics such as leaf area and plant structure; and climate variables are independent external drivers. Thus, the two findings do not contradict each other. Furthermore, the varying geometry between satellite sensors and the sun could also affect the signal of GPP/SIF across varying climates. However, difference between fluorescence acquisition due to different viewing zenith angles is much smaller than that between wet and dry climate zones (Biriukova et al., 2020), indicating that moisture control on the GPP/SIF pattern is not an artifact of sun-sensor geometry.

Since GPP/SIF is an integrator composite of GPP and SIF, mechanistic understanding of the observed spatial pattern of GPP/SIF may benefit from examining climatic controls of GPP and SIF separately. As Figure 2 shows, the spatial pattern of GPP resembles that of SIF, confirming findings of earlier studies (Frankenberg *et al*., 2011; Guanter *et al*., 2012; Joiner *et al*., 2013, Köhler *et al*., 2015b; Sun *et al*., 2017) and indicating that SIF can be a proxy of GPP at broad scales (Porcar-Castell *et al*., 2014; Guanter *et al*., 2014a). However, spatial concordance between GPP and SIF does not necessarily lead to constant GPP/SIF ratios, implying potentially asymmetric impacts of environmental variations on GPP and SIF. It is noteworthy that along the latitudinal gradient, while both GPP and SIF were unimodally distributed, GPP/SIF showed a bimodal distribution pattern. The high GPP/SIF ratio in the boreal region, despite low GPP or SIF, suggests that while temperature is arguably the primary factor shaping the latitudinal gradient of GPP or SIF, relatively high precipitation in the northern high latitude could contribute more to GPP than to SIF (Jeong *et al.*, 2012; Zhang *et al*., 2016c). Moreover, in analyzing climatic controls of global terrestrial productivity, Churkina and Running (1998) and Running *et al*. (2004) suggested that temperature is the primary limitation to productivity in high latitudes and solar radiation is the primary limitation in the tropics—two regional bands that showed highest GPP/SIF ratios. In contrast, when water is the primary limitation, the GPP/SIF ratio decreases. Since vapor-pressure deficit is often strongly related to precipitation (and influenced by temperature), stomatal response patterns could be key in explaining the observed bimodal latitudinal pattern of GPP/SIF ratios. In fact, in a broader climatic space, the GPP/SIF ratio decreased from wet-and-cold to dry-and-hot climate (Figures 3 and 4). This GPP/SIF pattern along a gradient of moisture availability is consistent with a key role of moisture controlled stomatal responses in determining the spatial variations of GPP/SIF ratios. When water is sufficient, stomata would be fully open and it is likely that other factors (primarily temperature and solar radiation) determine productivity and fluorescence emission symmetrically (Gerber & Häder, 1995; Guanter *et al*., 2010). However, recent studies suggest that under moisture deficits, stomatal closure could have asymmetrically larger impacts on GPP than on SIF (Reich *et al.*, 2018; Stocker *et al.*, 2019). This asymmetric impact of stomatal responses on GPP and SIF indicates that using SIF to approximate GPP needs to be carefully evaluated under drought versus normal conditions.

Furthermore, Gu *et al*. (2019) recently developed a novel photosynthesis model, in which GPP can be derived from SIF and a few physiological and structural parameters (see Eq. 23 in Gu *et al*. (2019)). In the Gu model, the parameter C_c that describes the CO_2 partial pressure inside the chloroplast stroma, and the parameter q_L representing the fraction of open PSII reaction centers, can both be impacted by moisture availability (Gu *et al*., 2019). Under moisture-stress, reduced stomatal conductance can lead to a decrease in both C_c and q_L, consequently resulting in reduced GPP/SIF ratio.

Several other parameters in the Gu model are worth noting. For instance, the fluorescence escaping probability ε is negatively related with GPP/SIF. ε is associated with canopy structure, with more complex canopies having a smaller fraction of escaped fluorescence (Fournier *et al.*, 2012). Therefore, a higher GPP/SIF may also imply more complex canopy structures with more layers. This canopy complexity can explain the high GPP/SIF in tropical rainforests. For northern boreal forests, the high GPP/SIF ratio may be explained by leaf shape. A significant proportion of northern boreal forests consist of needleleaf trees which may have a lower fluorescence escaping ratio due to their needle leaf shape and clustered leaf arrangement. Interestingly, GPP/SIF ratios were consistently high in the boreal region, although both GPP and SIF showed relatively low values in the northern boreal region. Many needle leaf species, such as pines, have round leaf shapes. Compared to flat broadleaves for which re-emitted fluorescence generally points to similar directions (canopy direction), these round leaves can re-emit fluorescence toward all directions (Van Wittenberghe *et al.*, 2015). Therefore, for the same amount of excited fluorescence, the satellite-captured escaped proportion is likely smaller for needleleaf forests than for broadleaf forests, leading to the higher GPP/SIF ratios in needleleaf forests for the same amount of GPP production. Field experiments comparing fluorescence escaping proportion across different vegetation types are needed to verify this hypothesis.

Another key parameter is the ratio of energy for thermal dissipation to that for fluorescence emission, k_{DF}. Based on the Gu model, a high fluorescence yield (i.e., a low k_{DF}) will lead to low GPP/SIF ratio. However, little information is available on how fluorescence yield may change among different vegetation types and across climatic gradients (Guanter *et al.*, 2014b). More field and experimental studies are needed to fill this gap on the variation of fluorescence yield. In particular, there is a need to explore the responses of different photosynthetic processes (i.e., photochemical carbon assimilation (GPP), fluorescence emission (SIF), and non-photochemical quenching (NPQ)) to

environmental stresses such as droughts (e.g., Huang *et al*., 2019; Piao *et al*., 2017; Sun *et al*., 2015) and shade (e.g., Chen *et al*., 2014; Gamon & Bond, 2013). A few studies have suggested that stresses increase NPQ but reduce GPP more than SIF, leading to reduced GPP/SIF ratios (Hilker *et al.*, 2008; Gu *et al.*, 2019).

In summary, our research based on spatially continuous data for both GPP and SIF evaluated the spatial and climatic distributions of GPP/SIF ratios and revealed surprising and important patterns. These findings, particularly the decreasing trend of GPP/SIF along a moisture gradient from cold-and-wet to hot-and-dry environments, are consistently corroborated with different GPP and SIF products. While the underlying mechanisms are still largely unknown and further studies are required, we hypothesize that from a climatic perspective, moisture-regulated stomatal responses are likely the key mechanism determining the spatial variation of GPP/SIF ratios. Furthermore, the high R^2 values (0.65-0.87) from machine learning algorithms used to model GPP/SIF by temperature, precipitation and solar radiation implied that the GPP/SIF ratio can be reliably predicted with these three key climatic variables, especially precipitation. Such a high predictability of GPP/SIF provides a useful approach to estimate GPP from SIF, and may increase the usefulness of SIF in constraining land surface model predictions of vegetation productivity.

Acknowledgements

The work is supported by an Oak Ridge National Lab subcontract (4000167205). J.M., D.R. and L.G. were supported by the Terrestrial Ecosystem Science Scientific Focus Area (TES SFA) project funded by the US Department of Energy, Office of Science, Office of Biological and Environmental Research. P.T. was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. Oak Ridge National Laboratory is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC05-00OR22725. J. X. and X. L. were supported by the National Aeronautics and Space Administration (NASA)'s Climate Indicators and Data Products for Future

National Climate Assessments (Grant number: NNX16AG61G). We also thank two anonymous reviewers for their constructive comments which helped improve our manuscript.

Data Availability Statement

The data that support the findings of this study are openly available, hosted by different agencies and institutions. The GOSAT SIF data are available on request from Christian Frankenberg (cfranken@caltech.edu). The GOME-2 SIF data are available at https://avdc.gsfc.nasa.gov/pub/data/satellite/MetOp/GOME_F/v28/. The GOSIF data are available at http://globalecology.unh.edu/. The TROPOMI SIF data are available at ftp://fluo.gps.caltech.edu/data/tropomi/. The FLUXCOM GPP data are available at https://www.bgc-jena.mpg.de/geodb/projects/Data.php. The MODIS GPP data are available at https://e4ftl01.cr.usgs.gov/MOLT/MOD17A2H.006/. The CRU climatic data are available at http://www. cru.uea.ac.uk/data/. The SMOS-IC soil moisture data are available at ftp://ext-catds-cecsm:catds2010@ftp.ifremer.fr/Land_products.

References

Al Bitar, A., Mialon, A., Kerr, Y. H., Cabot, F., Richaume, P., Jacquette, E., ... & Al-Yaari, A. (2017). The global SMOS Level 3 daily soil moisture and brightness temperature maps. *Earth System Science Data*, **9**, 293-315.

Biriukova, K., Celesti, M., Evdokimov, A., Pacheco-Labrador, J., Julitta, T., Migliavacca, M., ... & Rossini, M. (2020). Effects of varying solar-view geometry and canopy structure on solar-induced chlorophyll fluorescence and PRI. *International Journal of Applied Earth Observation and Geoinformation*, **89**, 102069.

Breiman, L. (2001). Random forests. *Machine Learning*, **45**, 5–32.

Campbell, P. K. E., Middleton, E. M., McMurtrey, J. E., Chappelle, E. W., *et al.* (2007). Assessment of vegetation stress using reflectance or fluorescence measurements. *Journal of Environmental Quality*, **36**, 832–845.

Chen, A., Lichstein, J. W., Osnas, J. L., & Pacala, S. W. (2014). Species-independent down-regulation of leaf photosynthesis and respiration in response to shading: evidence from six temperate tree species. *PloS one*, **9**(4), e91798.

Chen, A., Peng, S., & Fei, S. (2019). Mapping global forest biomass and its changes over the first decade of the 21st century. *Science China Earth Sciences*, **62**, 585-594.

Churkina, G., & Running, S. W. (1998). Contrasting climatic controls on the estimated productivity of global terrestrial biomes. *Ecosystems*, **1**, 206-215.

Coops et al., (2007). Comparison of MODIS, eddy covariance determined and physiologically modelled gross primary production (GPP) in a Douglas-fir forest stand. *Remote Sensing of Environment*, **107**, 385-401.

Damm, A., Elbers, J., Erler, A., Gioli, B., Hamdi, K., Hutjes, R., *et al.* (2010). Remote sensing of Sun-induced fluorescence to improve modeling of diurnal courses of gross primary production. *Global Change Biology*, **16**, 171–186.

Damm, A., Guanter, L., Paul-Limoges, E., Van der Tol, C., *et al.* (2015). Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches. *Remote Sens. Environ.* **166**, 91–105.

Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., ... & Wigneron, J. P. (2017). SMOS-IC: An alternative SMOS soil moisture and vegetation optical depth product. *Remote Sensing*, **9**, 457.

Frankenberg, C., & Berry, J. (2018). Solar induced chlorophyll fluorescence: origins, relation to photosynthesis and retrieval. In: *Reference Module in Earth Systems and Environmental Sciences: Comprehensive Remote Sensing*, Elsevier, Oxford, pp. 143-162

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. E., *et al*. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, **38**, L17706.

Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., Pollock, R., Taylor, T.E. (2014). Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2. *Remote Sens. Environ*. **147**, 1–12.

Frankenberg, C., Köhler, P., Magney, T. S., Geier, S., Lawson, P., Schwochert, M., *et al.* (2018). The Chlorophyll Fluorescence Imaging Spectrometer (CFIS), mapping far red fluorescence from aircraft. *Remote sensing of environment*, **217**, 523-536.

Fournier A, Daumard F, Champagne S, Ounis A, Goulas Y, Moya I. (2012). Effect of canopy structure on sun-induced chlorophyll fluorescence. *Journal of Photogrammetry and Remote Sensing*, **68**, 112–120.

Gamon, J. A., & Bond, B. (2013). Effects of irradiance and photosynthetic downregulation on the photochemical reflectance index in Douglas-fir and ponderosa pine. *Remote Sensing of Environment*, **135**, 141-149.

Gerber, S., & Häder, D. P. (1995). Effects of enhanced solar irradiation on chlorophyll fluorescence and photosynthetic oxygen production of five species of phytoplankton. *FEMS Microbiology Ecology*, **16**, 33-41.

Goulden, M.L.*et al.*(1996). Measurements of carbon sequestration by long term eddy covariance: Methods and a critical evaluation of accuracy. *Global Change Biology*, **2**, 169-182.

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing,* **27**, 659-678.

Gu, L., *et al*. (2019). Sun-induced Chl fluorescence and its importance for biophysical modeling of photosynthesis based on light reactions. *New Phytologist*, **223**,1179-1191.

Guan, K., *et al*. (2016). Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence. *Global change biology*, **22**, 716-726.

Guanter, L., Alonso, L., Gomez-Chova, L., Meroni, M., Preusker, R., Fischer, J., et al. (2010). Developments for vegetation fluorescence retrieval from spaceborne high-resolution spectrometry in the O_2-A and O_2-B absorption bands. *Journal of Geophysical Research-Atmospheres*, **115**.

Guanter, L. *et al.* (2012). Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. *Remote Sens. Environ.* **121**, 236–251.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., *et al.* (2014a). Linking large-scale chlorophyll fluorescence observations with cropland gross primary production. *Proceedings of the National Academy of Sciences*, **111**, E2511–E2511.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., *et al*. (2014b). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences*, **111**, E1327-E1333.

Gutman, G. G., Tarpley, D., and Ohrtng, G. et al. (1987), Cloud screening for determination of land surface characteristics in a reduced resolution satellite data set. *International Journal of Remote Sensing*, 8, 859-870.

Harris, I. P. D. J., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *International journal of climatology*, **34**, 623-642.

Hilker, T., Coops, N. C., Wulder, M. A., Black, T. A., & Guy, R. D. (2008). The use of remote sensing in light use efficiency based models of gross primary production: A review of current status and future requirements. *Science of the Total Environment*, **404**, 411-423.

Huang, M., Piao, S., Ciais, P., Peñuelas, J., Wang, X., Keenan, T. F., . . . Mao, J. (2019). Air temperature optima of vegetation productivity across global biomes. *Nature Ecology & Evolution*, 1. doi:10.1038/s41559-019-0838-x

Huemmrich, K. F., Campbell, P., Landis, D., & Middleton, E. (2019). Developing a common globally applicable method for optical remote sensing of ecosystem light use efficiency. *Remote Sensing of Environment*, 230, 111190.

Jackson, T. J., Bindlish, R., Cosh, M. H., Zhao, T., Starks, P. J., Bosch, D. D., ... & Leroux, D. (2011). Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the US. *IEEE Transactions on Geoscience and Remote Sensing*, **50**, 1530-1543.

Jeong, S.J, C.H. Ho, B.M. Kim, S. Feng, and D. Medvigy, 2012, Nonlinear response of vegetation to coherent warming over northern high latitudes. *Remote Sensing Letters*, doi:10.1080/2150740X.2012.703790.

Jeong, S.J, D. Schimel, C. Frankenberg, D.T. Drewry, J.B. Fisher, M. Verma, J.A. Berry, J.E. Lee, and J. Joiner, 2017, Application of satellite solar-induced chlorophyll fluorescence to understanding large-scale variations in vegetation phenology and function over northern high latitude forests. *Remote sensing of Environment*, 190, 178-187.

Joiner, J., Yoshida, Y., Vasilkov, A., Middleton, E. (2011). First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences*, **8**, 637–651.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. *et al.* (2013). Global monitoring of terrestrial chlorophyll fluorescence from moderate spectral resolution near-infrared satellite measurements: Methodology, simulations, and application to GOME-2. *Atmos. Meas. Tech.*, **6**, 2803–2823.

Jung, M., *et al*. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research: Biogeosciences*, **116**.

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., . . . Friedlingstein, P. (2017). Compensatory water effects link yearly global land CO2 sink changes to temperature. *Nature*, **541**, 516.

Kaufman, Y. J., Holben, B. N.*et al.* (1993). Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint, and desert reflection. *International Journal of Remote Sensing*, **14**, 21-52.

Köhler, P., Guanter, L., & Frankenberg, C. (2015a). Simplified physically based retrieval of sun-induced chlorophyll fluorescence from GOSAT data. *IEEE Geoscience and Remote Sensing Letters*, **12**, 1446-1450.

Köhler, P., Guanter, L., Joiner, J. (2015b). A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data. *Atmos. Meas. Tech.* **8**, 2589–2608.

Köhler P, Frankenberg C, Magney T S, *et al*. (2018). Global retrievals of solar-induced chlorophyll fluorescence with TROPOMI: First results and intersensor comparison to OCO-2. *Geophysical Research Letters*, **45**, 10456-10463.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., ... & Wohlfahrt, G. (2010). Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. *Global Change Biology*, *16*(1), 187-208.

Lee, J.-E., Berry, J. A., Tol, C., Yang, X., Guanter, L., Damm, A., Baker, I., & Frankenberg, C. (2015). Simulations of chlorophyll fluorescence incorporated into the community land model version 4. *Global Change Biology*, **21**, 3469–3477.

Li, X., Xiao, J. F., He, B. B., Arain, M. A., Beringer, J., Desai, A. R., … Varlagin, A. (2018a). Solar-induced chlorophyll fluorescence is strongly correlated with terrestrial photosynthesis for a wide variety of biomes: First global analysis based on OCO-2 and flux tower observations. *Global Change Biology*, **24**, 3990–4008.

Li, X., Xiao, J. F., & He, B. B. (2018b). Chlorophyll fluorescence observed by OCO-2 is strongly related to gross primary productivity estimated from flux towers in temperate forests. *Remote Sensing of Environment*, **204**, 659–671.

Li, X., & Xiao, J. (2019). A Global, 0.05-Degree Product of Solar-Induced Chlorophyll Fluorescence Derived from OCO-2, MODIS, and Reanalysis Data. *Remote Sensing*, **11**, 517.

MacBean, N., *et al*. (2018) Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. *Scientific reports*, **8**, 1973.

Mao, J., Ribes, A., Yan, B., Shi, X., Thornton, P. E., Séférian, R., ... & Zhu, Z. (2016). Human-induced greening of the northern extratropical land surface. *Nature Climate Change*, **6**(10), 959-963.

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., ... & Running, S. W. (2003). Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, **300**, 1560-1563.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Koppen-Geiger climate classification. *Hydrol. Earth Syst. Sci*, **11**, 1633-1644.

Parazoo, N. C., Bowman, K., Fisher, J. B., Frankenberg, C., Jones, D. B., Cescatti, A., ... & Montagnani, L. (2014). Terrestrial gross primary production inferred from satellite fluorescence and vegetation models. *Global change biology*, **20**(10), 3103-3121.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., . . . Zeng, N. (2013). Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends. *Global Change Biology*, **19**, 2117-2132.

Piao, S., Liu, Z., Wang, T., Peng, S., Ciais, P., Huang, M., . . . Tans, P. P. (2017). Weakening

temperature control on the interannual variations of spring carbon uptake across northern lands. *Nature Clim. Change*, 7(5), 359-363.

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., ... & Nemani, R. R. (2020). Characteristics, drivers and feedbacks of global greening. *Nature Reviews Earth & Environment*, **1**, 14-27.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., Van der Tol, C., Flexas, J., Pfündel, E.E., et al. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges. *Journal of Experimental Botany*, **65**, 4065–4095.

Qiu, B., Li, W., Wang, X., Shang, L., Song, C., Guo, W., & Zhang, Y. (2019). Satellite-observed solar-induced chlorophyll fluorescence reveals higher sensitivity of alpine ecosystems to snow cover on the tibetan plateau. *Agricultural and Forest Meteorology*, **271**, 126-134.

Reich, P. B., Sendall, K. M., Stefanski, A., Rich, R. L., Hobbie, S. E., & Montgomery, R. A. (2018). Effects of climate warming on photosynthesis in boreal tree species depend on soil moisture. *Nature*, **562**, 263.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, **566**, 195-204.

Rossini, M., Nedbal, L., Guanter, L., Ac, A., Alonso, L., Burkart, A., Cogliati, S., Colombo, R., Damm, A., Drusch, M., Hanus, J., Janoutova, R., Julitta, T., Kokkalis, P., Moreno, J.,Novotny, J., Panigada, C., Pinto, F., Schickling, A., SchÃijttemeyer, D., Zemek, F., & Rascher, U. (2015). Red and far red sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geophysical Research Letters*, **42**, 1632–1639.

Rossini, M., Meroni, M., Celesti, M., Cogliati, S., Julitta, T., Panigada, C., Rascher, U., Tol, C.v.d., & Colombo, R. (2016). Analysis of red and far-red sun-induced chlorophyll fluorescence and their ratio in different canopies based on observed and modeled data. *Remote Sensing*, **8**, 412.

Running, S. W., Thornton, P. E., Nemani, R., & Glassy, J. M. (2000). Global terrestrial gross and net primary productivity from the earth observing system. In *Methods in Ecosystem Science*, Springer, New York, pp. 44-57.

Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M.S., Reeves, M., Hashimoto, H., 2004. A continuous satellite-derived measure of global terrestrial primary production. *Bioscience* 54,

547–560.

Schimel, D., Pavlick, R., Fisher, J. B., Asner, G. P., Saatchi, S., Townsend, P., ... & Cox, P. (2015). Observing terrestrial ecosystems and the carbon cycle from space. *Global Change Biology*, *21*(5), 1762-1776.

Smith, W.K., Biederman, J.A., Scott, R.L., Moore, D.J.P., He, M., Kimball, J.S., Yan, D., Hudson, A., Barnes, M.L., MacBean, N. *et al*. (2018). Chlorophyll fluorescence better captures seasonal and interannual gross primary productivity dynamics across dryland ecosystems of southwestern North America. *Geophysical Research Letters*, **45**, 748–757.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., & Peñuelas, J. (2019). Drought impacts on terrestrial primary production underestimated by satellite monitoring. *Nature Geoscience*, **12**, 264.

Sun, Y., Fu, R., Dickinson, R., Joiner, J., Frankenberg, C., Gu, L., ... & Fernando, N. (2015). Drought onset mechanisms revealed by satellite solar-induced chlorophyll fluorescence: Insights from two contrasting extreme events. *Journal of Geophysical Research: Biogeosciences*, **120**(11), 2427-2440.

Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D. S., Jung, M., Guanter, L., … Yuen, K. (2017). OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. *Science*, **358,** eaam5747.

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Köhler, P., Magney, T. (2018). Overview of Solar-Induced Chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, Cross-Mission Comparison, and Global Monitoring for GPP. *Remote Sensing of Environment*, **209**, 808-823.

Tan, C., *et al*. (2013) Using hyperspectral vegetation indices to estimate the fraction of photosynthetically active radiation absorbed by corn canopies. *International journal of remote sensing*, **34**, 8789-8802.

Thum, T., *et al*. (2017). Modelling sun-induced fluorescence and photosynthesis with a land surface model at local and regional scales in northern Europe. *Biogeosciences*, **14**, 1969-1987.

Tramontana, G., Jung, M., *et al*. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, *Biogeosciences*, **13**, 4291-4313.

Tucker, C. J., Fung, I. Y., Keeling, C. D., & Gammon, R. H. (1986). Relationship between atmospheric CO 2 variations and a satellite-derived vegetation index. *Nature*, **319**, 195-199.

Van Wittenberghe, S., *et al*. (2015). Bidirectional sun-induced chlorophyll fluorescence emission is influenced by leaf structure and light scattering properties—A bottom-up approach. *Remote Sensing of Environment* ,**158**, 169-179.

Verrelst, J., *et al.* (2015). Global sensitivity analysis of the scope model: What drives simulated canopy-leaving sun-induced fluorescence? *Remote Sensing of Environment*, **166**,8-21.

Verrelst, J., Tol, C.v.d., Magnani, F., *et al.* (2016). Evaluating the predictive power of sun-induced chlorophyll fluorescence to estimate net photosynthesis of vegetation canopies: A scope modeling study. *Remote Sensing of Environment*, **176**, 139–151.

Viovy, N., 2016. CRUNCEP data set.

Wei, S., Yi, C., Fang, W., & Hendrey, G. (2017). A global study of GPP focusing on light-use efficiency in a random forest regression model. *Ecosphere*, **8**, e01724.

Xiao, J.F., Zhuang, Q., Law, B.E., Chen, J., Baldocchi, D.D., Cook, D.R., Oren, R., Richardson, A.D., Wharton, S., Ma, S., Martin, T.A., Verma, S.B., Suyker, A.E., Scott, R.L., Monson, R.K., Litvak, M., Hollinger, D.Y., Sun, G., Davis, K.J., Bolstad, P.V., Burns, S.P., Curtis, P.S., Drake, B.G., Falk, M., Fischer, M.L., Foster, D.R., Gu, L., Hadley, J.L., Katul, G.G., Matamala, R., McNulty, S., Meyers, T.P., Munger, J.W., Noormets, A., Oechel, W.C., Paw U, K.T., Schmid, H.P., Starr, G., Torn, M.S., and Wofsy, S.C. (2010) A continuous measure of gross primary production for the conterminous United States derived from MODIS and AmeriFlux data. *Remote Sensing of Environment*, 114 (3), 576-591.

Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J.A., Huete, A.R., Ichii, K., Ni, W., Pang, Y., Rahman, F.F., Sun, G., Yuan, W., Zhang, L., Zhang, X. (2019a) Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sensing of Environment*, 233, 111383. https://doi.org/10.1016/j.rse.2019.111383.

Xiao, J., *et al*. (2019b) Solar-induced chlorophyll fluorescence exhibits a universal relationship with gross primary productivity across a wide variety of biomes. *Global change biology*, **25**, e4-e6.

Yang, H., Yang, X., Zhang, Y., *et al.* (2017) Chlorophyll fluorescence tracks seasonal variations of photosynthesis from leaf to canopy in a temperate forest. *Glob. Chang. Biol.*, **23**, 2874–2886.

Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., William Munger, J., Kornfeld, A., & Richardson, A. D. (2015). Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. *Geophysical Research Letters*, **42**, 2977–2987

Zhang, Y., Song, C., Sun, G., Band, L. E., Noormets, A., & Zhang, Q. (2015). Understanding moisture stress on light use efficiency across terrestrial ecosystems based on global flux and remote-sensing data. *Journal of Geophysical Research: Biogeosciences*, **120**, 2053-2066.

Zhang, Y., *et al*. (2016a). Consistency between sun-induced chlorophyll fluorescence and gross primary production of vegetation in North America. *Remote Sensing of Environment*, **183**, 154-169.

Zhang, Y., Guanter, L., Berry, J. A., Tol, C.v.d., Yang, X., Tang, J., & Zhang, F. (2016b). Model-based analysis of the relationship between sun-induced chlorophyll fluorescence and gross primary production for remote sensing applications. *Remote Sensing of Environment*, **187**, 145–155.

Zhang, Y., Xiao, X., Guanter, L., Zhou, S., Ciais, P., Joiner, J., ... & Kato, E. (2016c). Precipitation and carbon-water coupling jointly control the interannual variability of global land gross primary production. *Scientific reports*, **6**, 39748.

Zhang, Z. Y., Zhang, Y. G., Joiner, J., & Migliavacca, M. (2018). Angle matters: Bidirectional effects impact the slope of relationship between gross primary productivity and sun-induced chlorophyll fluorescence from Orbiting Carbon Observatory-2 across biomes. *Global Change Biology*, **24**, 5017–5020.

Zhao, M., & Running, S. (2010). Drought-induced reduction in global terrestrial net primary production from 2000 through 2009, *Science*, **329**, 940–943.

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., ... & Cao, C. (2016). Greening of the Earth and its drivers. *Nature climate change*, **6**, 791-795.

Figure Legends

Figure 1. Growing season GPP/SIF ratio and its temporal variation over the global land. (a) Spatial distribution of averaged growing season GPP/SIF ratio derived from multiple gridded SIF and GPP products. Note that both GPP and SIF are normalized before calculating GPP/SIF ratios. Left panel shows the change of averaged GPP/SIF values along the latitudinal gradient. (b) Spatial distribution of the standard deviation of the ensemble of growing season GPP/SIF ratio. The GPP/SIF ratios are calculated as mean GPP from MODIS and FLUXCOM divided by the four SIF products. Similarly, the left panel is the change of averaged GPP/SIF values along the latitudinal gradient.

Figure 2. Spatial (**a, c**) and climatic (**b, d**) distributions of averaged growing season GPP (**a, b**) and SIF (**c, d**). Note both GPP and SIF are normalized before calculating GPP/SIF ratios. Left panels on **a** and **c** shows the change of averaged GPP or SIF values along the latitudinal gradient. (**e**) The relationship between GPP/SIF and GPP. (**f**) The relationship between GPP/SIF and SIF. The curve fitting in panels e and f is based on a quadratic polynomial model.

Figure 3. The distribution of the GPP/SIF ratio in the temperature-precipitation space (a) and in the APAR-precipitation space (b). Note that both GPP and SIF are normalized before calculating GPP/SIF ratios. We used climate data of the growing season (GS). Each bin is 4°C (temperature) by 240mm (precipitation), or 5 W m^{-2} (APAR) by 240mm (precipitation).

Figure 4. The distribution of monthly GPP/SIF during the growing season in the climatic space derived from flux tower GPP and corresponding satellite SIF products. Note that both GPP and SIF are normalized before calculating GPP/SIF ratios. Each bin is 4°C (temperature) by 20mm (precipitation).

Figure 5. Predicted vs observed GPP/SIF by the Random Forest machine learning, for (a) global, (b) boreal, (c) temperate and (d) tropical regions, respectively. Here independent variables include shortwave radiation (SWR), temperature (TMP) and precipitation (PRE).

Figure 1

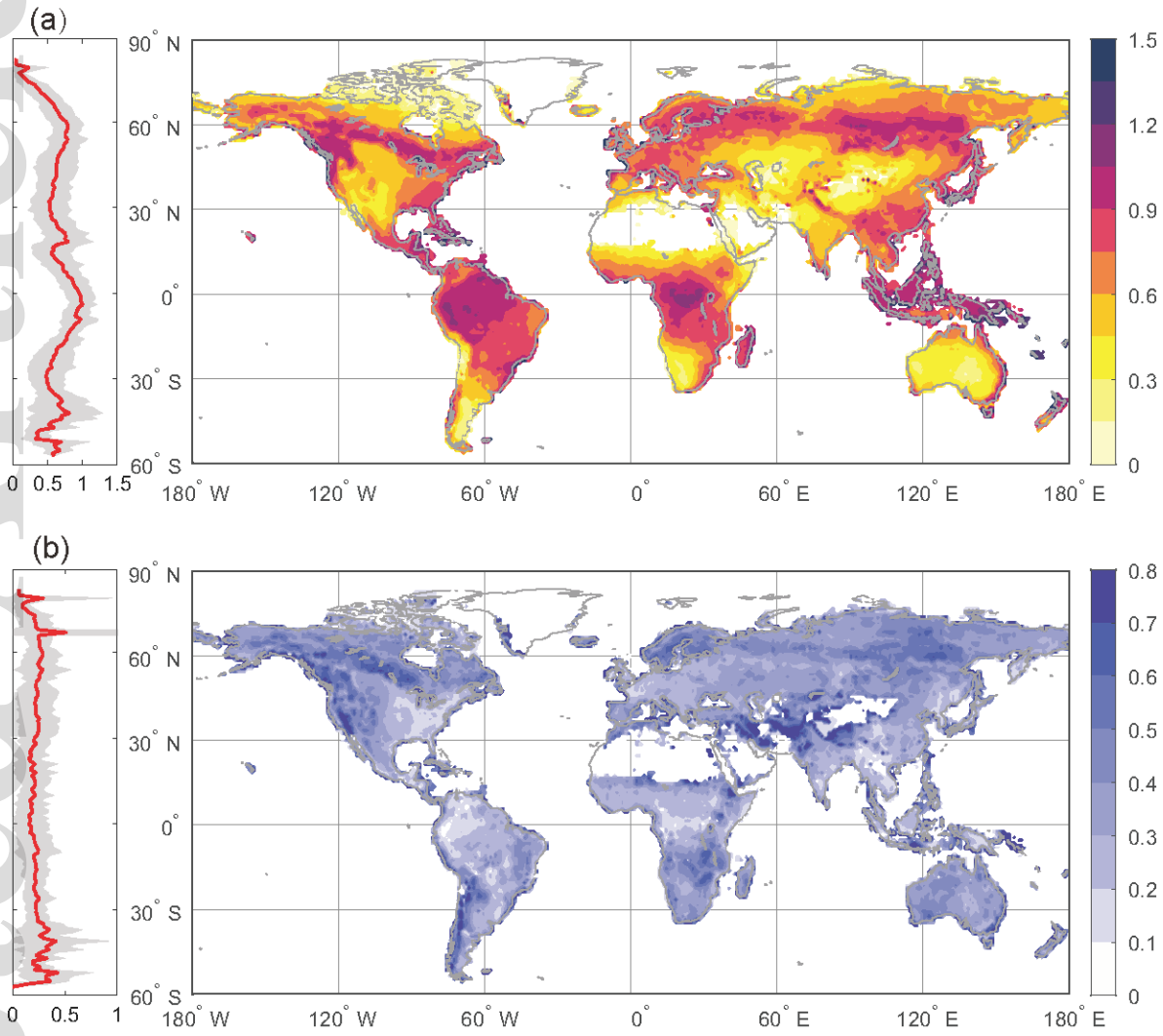


Figure 2

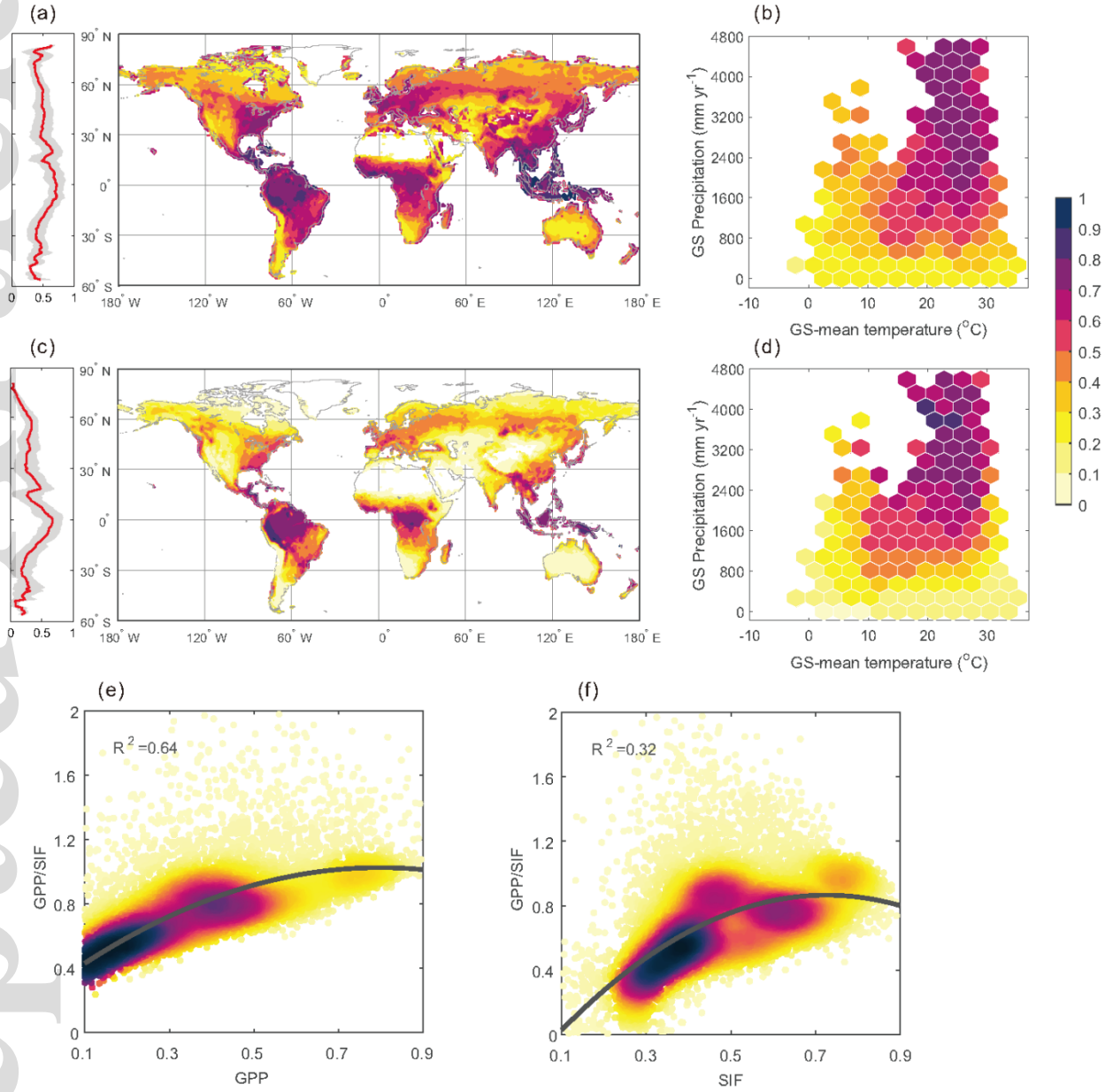


Figure 3

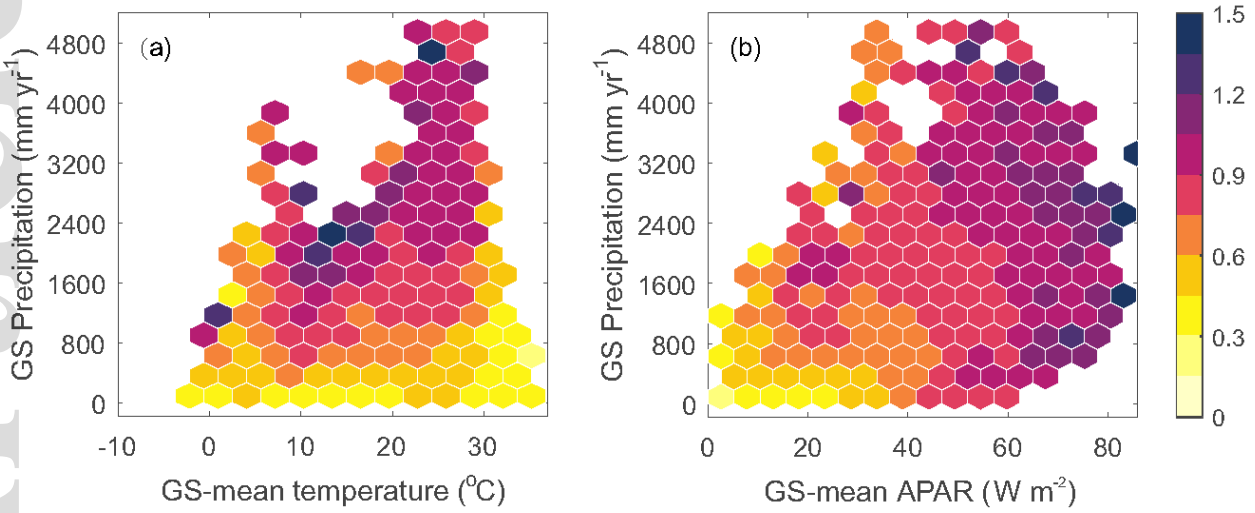


Figure 4

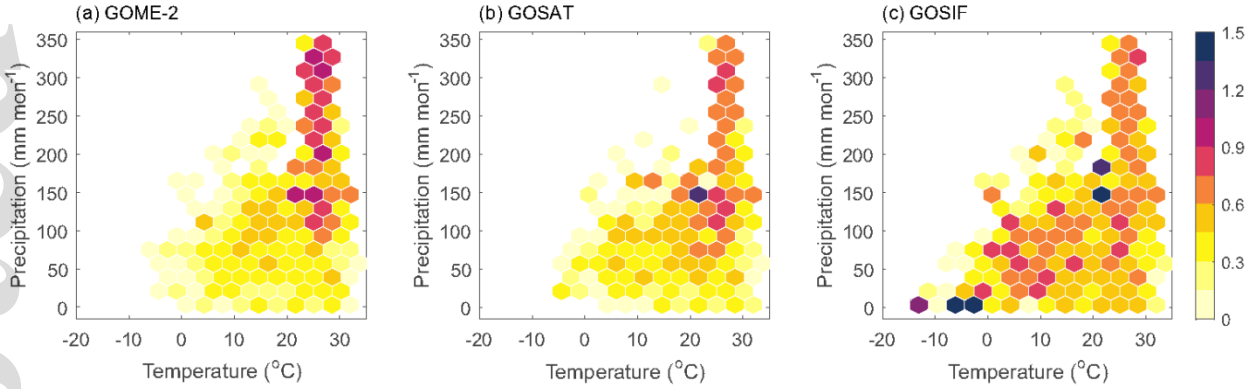
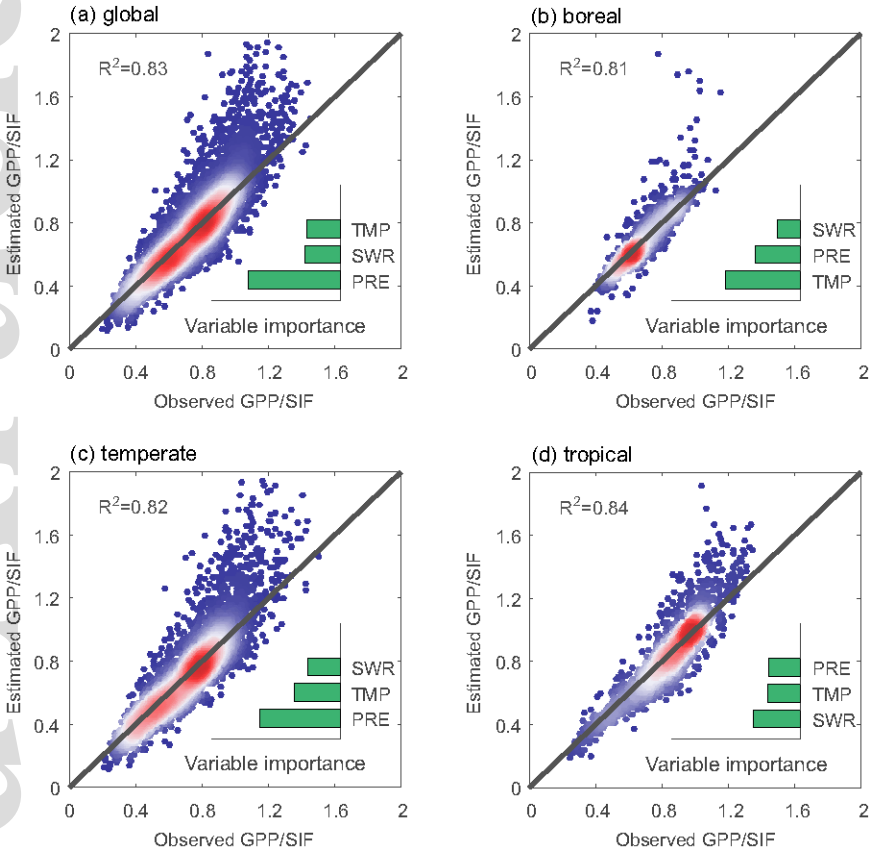
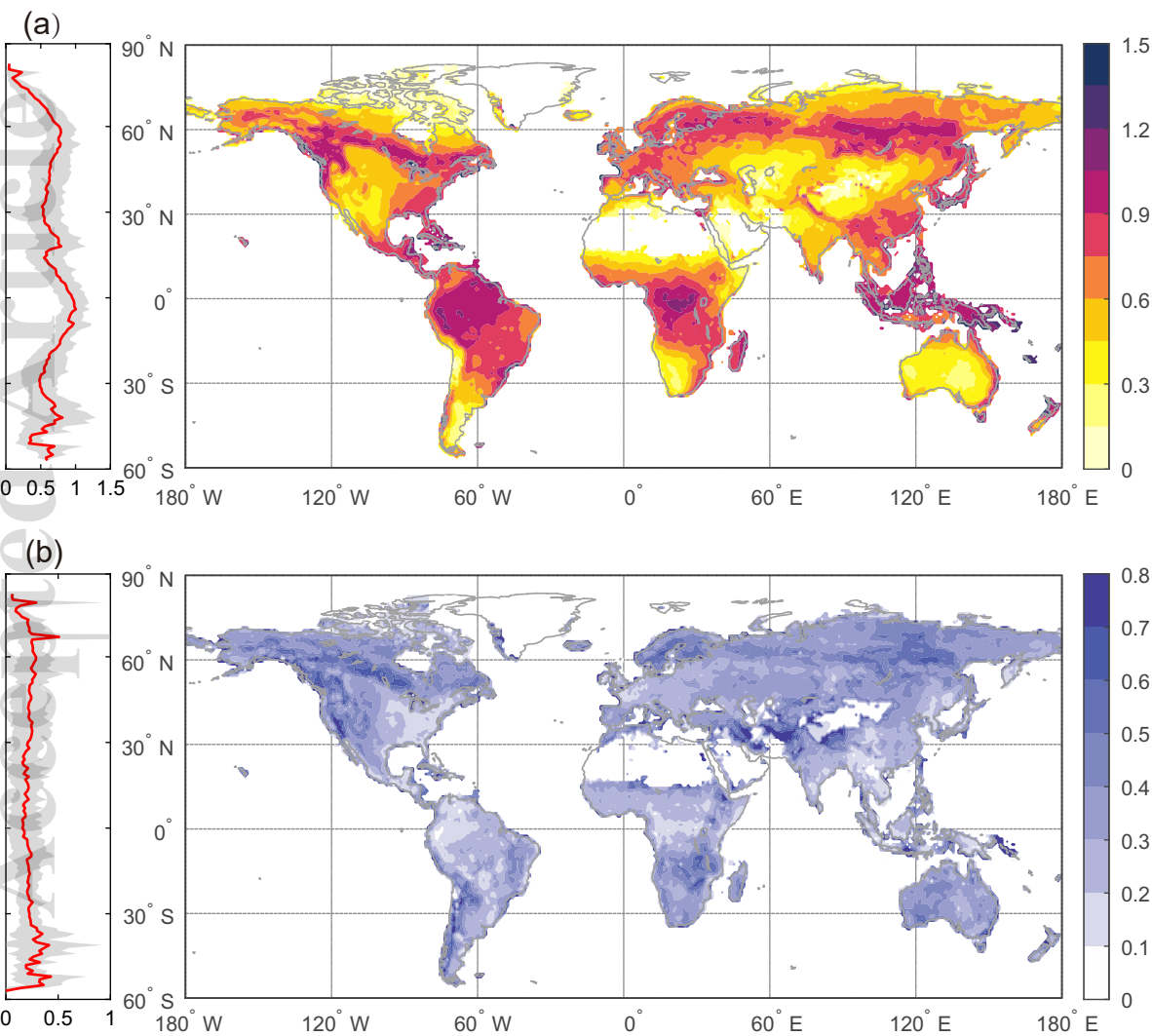


Figure 5




(a)
90° N
60° N
30° N
0°
30° S
60° S
180° W
120° W
60° W
0°
60° E
120° E
180° E
0
0.5
1
1.5
0
0.3
0.6
0.9
1.2
1.5
(b)
90° N
60° N
30° N
0°
30° S
60° S
180° W
120° W
60° W
0°
60° E
120° E
180° E
0
0.5
1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8